

DOCUMENT RESUME

ED 174 673

TH 009 645

AUTHOR Powell, George; And Others
TITLE The Equipercentile Assumption as a Pseudo-Control Group Estimate of Gain.
PUB DATE Apr 79
NOTE 10p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Gains; Compensatory Education Programs; Elementary Education; *Evaluation Methods; *Expectation; Models; Norm Referenced Tests; *Post Testing; Pretests; *Program Evaluation; Scores; Validity

IDENTIFIERS Elementary Secondary Education Act Title I; *Equipercentile Assumption; Normal Curve Equivalent Scores; *RMC Models

ABSTRACT

The tenability of the equipercentile growth assumption was investigated. This assumption is the basis of the norm referenced evaluation model (Model A) for Title I program evaluations, and supposes that a cohort of students, given no special educational intervention, will maintain the same percentile rank over the course of a school year, as measured on a standardized achievement test. Model A also assumes that a good estimate of the no-treatment expectation is the percentile corresponding to the average pretest score. Data were analyzed from reading and mathematics scores on the Metropolitan Achievement Tests; pre- and post tests were administered to elementary school students in grades two through four. For reading test results, the equipercentile growth method of estimation was found to be valid within the limits of the sample size. The comparisons between pretest and post test means for the mathematics scores, however, were significantly different from what would be expected using the equipercentile assumption. (GDC)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED174673

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

The Equipercentile Assumption as a
Pseudo-Control Group Estimate
of Gain

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

George Powell

George Powell
Jerome Schmidt
Paul Raffeld

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Resource Development Institute, Inc.
Title I Evaluation Technical Assistance Center

April 1979

Paper presented at the Annual Meeting
of the American Educational Research
Association, San Francisco, California
April 8-13, 1979.

AERA SESSION # 15.05

IM009 645

The Equipercntile Assumption as a Psuedo-Control Group Estimate of Gain

The objective of this study was to investigate the tenability of the assumption that a cohort of students given no special educational intervention will, on the average, maintain the same percentile rank over the course of a school year as measured on a common standardized achievement test.

An underlying assumption of the widely used Title I evaluation and reporting system's norm-referenced evaluation model (Model A) is that a good estimate of no-treatment expectation is the percentile corresponding to the average pretest standard score.

Tallmadge and Wood (1976) state that "It is assumed that, without the Title I treatment, the status of the group at posttest time would be the same as it was at pretest time. Thus, the group's pretest percentile becomes the expected no-treatment posttest performance." (page 4) There has been no documented attempt to date to verify this equipercntile assumption. The problem of testing this assumption is compounded by the fact that method of selection into special treatment groups such as Title I will affect the measured percentile gain from pretest to posttest (Tallmadge, 1976). That is, a simple comparison of pretest and posttest percentiles for a given group of Title I children might differ not due to treatment, but due to the fact that the group membership was determined by low pretest scores. Conversely, for a group of Title I selected children in an ineffective program or for whom a program is withheld, the change from pretest to posttest percentile would not be assumed to be zero but rather some positive value reflecting regression of the group mean toward the population mean.

Since an overwhelming majority of school districts implementing the Title I evaluation and reporting system are choosing Model A and, thus

relying on the validity of its no-treatment expectation--the equipercentile assumption--its validity is of some concern to evaluators of Title I programs.

Method

Data were collected from a Texas school district with an average daily attendance of approximately 3,000 children grades 1-12. Data consisted of standard scores for Total reading and Total mathematics on the Metropolitan Achievement Test which was administered both during the Spring of 1977 and the Spring of 1978. Data were used only from students who had both pre and posttest scores available either on the Total reading or the Total math scales. Because of data availability constraints, three groups of pre-post test cycles were analyzed. These cycles were the end-of-second grade to end-of-third grade matched scores, the end-of-fourth and end-of-fifth grade matched scores, and a third group of matched scores consisting of an end-of-second (1976) to end of fourth (1978) grade cycle. Within each of these cycles, groups of matched scores were formed on the basis of the pre-test standard score. Each cycle was divided into approximate thirds forming a low, medium and high group within each cycle. Table I outlines the N-size and test instrument used in each group and cycle. It should be noted in Table I that N-size for total reading and total math scores within each cycle vary due to the presence of some students without matched scores on one of those two scales. Also to be noted on Table I is that there is a change in test level from pretest to posttest in each of the three cycles. The range of pretest scores was truncated at the 10th and 90th percentiles to avoid measurement problems in the extremes of the range.

Once the data were collected, means and standard deviations were computed for each group within each cycle as well as all scores within a cycle. All computations were done with standard scores. It was not until

Table I
Data Collection Design

	Cycle					
	Grade 2→Grade 3		Grade 4→Grade 5		Grade 2→Grade 4	
	Total Reading	Total Math	Total Reading	Total Math	Total Reading	Total Math
Percentile Boundaries for						
- Low Group	10-40	10-40	10-30	10-30	10-38	10-38
- Medium Group	44-60	44-60	32-60	32-50	44-56	44-58
- High Group	62-90	62-90	62-90	52-90	62-90	62-90
Total number of pretests	248	248	263	263	283	283
Matched pre posttests	148	148	183	183	136	136
Non Title I, matched between 10-90%ile (i.e., total N-size)	109	104	136	130	81	84
N for Low group	39	40	47	54	33	35
N for Medium group	34	22	44	38	23	20
N for High group	36	42	45	38	25	29
Level/Form Administered:						
Pretest	Primary II/F		Elementary/F		Primary II/F	
Posttest	Elementary/F		Intermediate/F		Elementary/F	

means for each group were found that conversion to percentiles was accomplished using norm tables provided in the MAT Teacher's Handbook.

Results

Table II lists the obtained mean standard scores and percentile equivalents for each group within each cycle. In addition, Table II lists the Normal Curve Equivalent (NCE) value corresponding to each percentile value. Since NCEs are standard scores with the same range as percentiles, the difference between pre and posttest NCEs provides a more linear comparison than do percentile pre and posttest values.

A visual inspection of Table II shows that the equipercentile assumption is the exception rather than the rule. The same visual inspection will show that NCE values for pretest and posttest means are within 1-1/2 NCE of each other for Total Reading scores for all three cycles. For Total Math scores, the Grade 4 to Grade 5 cycle also displays this near equivalence.

Of course, visual inspection with N-sizes in the 30-40 range can be overly conservative. The next step was to perform a t-test for correlated scores where the population means pretest and posttest are estimated by the midpoint of the range of percentile scores. For example for the Total N within a cycle, the standard score equivalent of the 50th percentile was the appropriate value for the population mean. Table III outlines the estimated population means, the obtained means, the prepost test correlation, and the standard error of the difference of means for all groups.

There are two major results depicted in Table III. First, none of the comparisons for Total Reading scores differ significantly from the population estimates of the means. Since the population estimates of means follow the same logic as the equipercentile assumption, we can state that for Total Reading scores, the equipercentile assumption is valid within the limits of

Table II
Summary of Results

Cycle Pretest/ Posttest Grade	Subject Matter	N	Group	Pretest SS	%ile	NCE	Posttest SS	%ile	NCE
2→3	Reading	(109)	Total	53.26	53.04	51.6	61.52	52.08	51.1
		(39)	10-40%ile	47.49	25.45	36.1	53.64	26.56	36.8
		(34)	44-60%ile	52.38	49.52	49.8	60.47	47.88	48.9
		(36)	62-90%ile	60.33	78.66	66.7	71.06	78.12	66.3
	Math	(104)	Total	58.45	55.80	53.1	67.76	39.52	44.4
		(40)	10-40%ile	52.40	25.60	36.2	61.48	21.44	33.3
		(22)	44-60%ile	57.32	51.28	50.7	65.09	32.18	40.3
		(42)	62-90%ile	64.81	77.62	66.0	75.14	66.28	58.8
4→5	Reading	(136)	Total	69.05	46.10	47.9	76.45	43.80	46.7
		(47)	10-30%ile	58.04	18.08	30.8	67.60	19.20	31.7
		(44)	32-60%ile	68.55	44.20	46.9	76.50	44.00	46.8
		(45)	62-90%ile	81.04	76.04	64.9	85.64	73.28	63.1
	Math	(130)	Total	78.58	39.16	44.2	84.21	38.84	44.0
		(54)	10-30%ile	71.41	21.64	33.5	78.48	20.96	33.0
		(38)	32-50%ile	79.34	40.68	45.0	85.74	44.96	47.3
		(38)	52-90%ile	88.03	70.12	61.1	90.82	65.28	58.3
2→4	Reading	(81)	Total	52.21	48.84	49.4	69.53	47.06	48.4
		(33)	10-38%ile	47.48	25.40	36.1	62.09	28.18	37.8
		(23)	44-56%ile	52.39	49.56	49.8	72.13	54.26	52.2
		(25)	62-90%ile	58.28	72.56	62.6	76.96	67.92	59.8
	Math	(84)	Total	58.04	54.16	52.2	77.73	36.92	43.0
		(35)	10-38%ile	52.69	26.76	37.0	72.09	24.18	35.2
		(20)	44-58%ile	57.00	50.00	50.0	78.95	39.90	44.6
		(29)	62-90%ile	65.21	78.42	66.6	83.69	54.76	52.5

Table III

Requisite information for computing a t-test for correlated means assuming known population means

Cycle/Group	Subject Matter	$\bar{X}_{pre}$	$\bar{X}_{post}$	μ_{pre}	μ_{post}	$r_{pre-post}$	$s_{\bar{X}_{pre-post}}$	t
Grade 2→3								
Total	Total Reading	53.26	61.52	52.5	61.0	0.79	0.55	0.4
Low		47.49	53.64	47.0	52.5	0.38	0.99	0.7
Medium		52.38	60.47	53.0	61.5	0.44	0.80	0.5
High		60.33	71.06	60.6	71.0	0.30	0.93	0.3
Total	Total Math	58.45	67.76	57.0	70.5	0.74	0.58	7.3*
Low		52.40	61.48	51.8	62.0	0.39	1.06	1.1
Medium		57.32	65.09	57.5	71.0	0.55	1.16	4.9*
High		64.81	75.14	65.0	80.3	0.29	0.81	6.2*
Grade 4→5								
Total	Total Reading	69.05	76.45	71.0	78.5	0.83	0.50	0.2
Low		58.04	67.70	58.0	67.0	0.26	0.93	0.6
Medium		68.55	76.50	68.8	76.8	0.55	0.66	0.1
High		81.04	85.64	82.5	88.0	0.70	0.87	1.0
Total	Total Math	78.58	84.21	82.0	87.0	0.72	0.49	1.3
Low		71.41	78.48	70.0	77.0	0.15	0.77	0.1
Medium		79.34	85.74	79.2	84.7	0.29	0.60	1.5
High		88.03	90.82	89.2	94.2	0.28	0.84	2.6*
Grade 2→4								
Total	Total Reading	52.21	69.53	52.5	71.0	0.67	0.92	1.3
Low		47.48	62.09	47.0	60.5	0.46	1.19	0.9
Medium		52.39	72.13	53.0	71.5	0.12	1.65	0.8
High		58.28	76.96	60.0	82.3	0.47	1.86	1.9
Total	Total Math	58.04	77.73	57.0	82.0	0.50	0.90	5.9*
Low		52.69	72.09	51.8	71.8	0.31	1.10	0.5
Medium		57.00	78.95	57.5	83.0	-0.08	1.42	2.5*
High		65.21	83.69	65.0	90.3	-0.09	2.02	3.4*

*p < .05

the sample size. The second major result is that the majority of comparisons among the pretest and posttest means for the Total Math Scores are significantly different from what we would expect following the logic of the equipercentile assumption.

Discussion

The results presented here are intriguing but hardly compelling. What is of interest, however, is the need for a paradigm such as the one outlined above, that will enable a local school district to determine if the equipercentile assumption does hold for its current sample of children. Of even greater interest would be the possibility of deriving a locally valid alternative assumption that would describe achievement gain under the hypothesis of no-treatment.

For the school district whose data we used, I would feel much more confident in relating Title I gains to the no-treatment expectation provided by the equipercentile assumption for a Title I reading program than I would a Title I math program.

There are two methodological flaws that we are currently overcoming in additional local school data analysis. The first is the changing of levels between pretest and posttest. The result of this change is unknown vis-a-vis the equipercentile assumption. The second flaw was the decision to truncate the available pretest scores below the 10th percentile and above the 90th. Especially for Title I project evaluation, more information is needed for the behavior of the equipercentile assumption in the extreme low range of the distribution.

References

Tallmadge, G. K., and Wood, C. T. User's Guide, ESEA Title I Evaluation and Reporting System. RMC Research Corporation, 1976.

Tallmadge, G. K. The Regression Effect, Technical Paper No. 3. RMC Research Corporation, 1976.